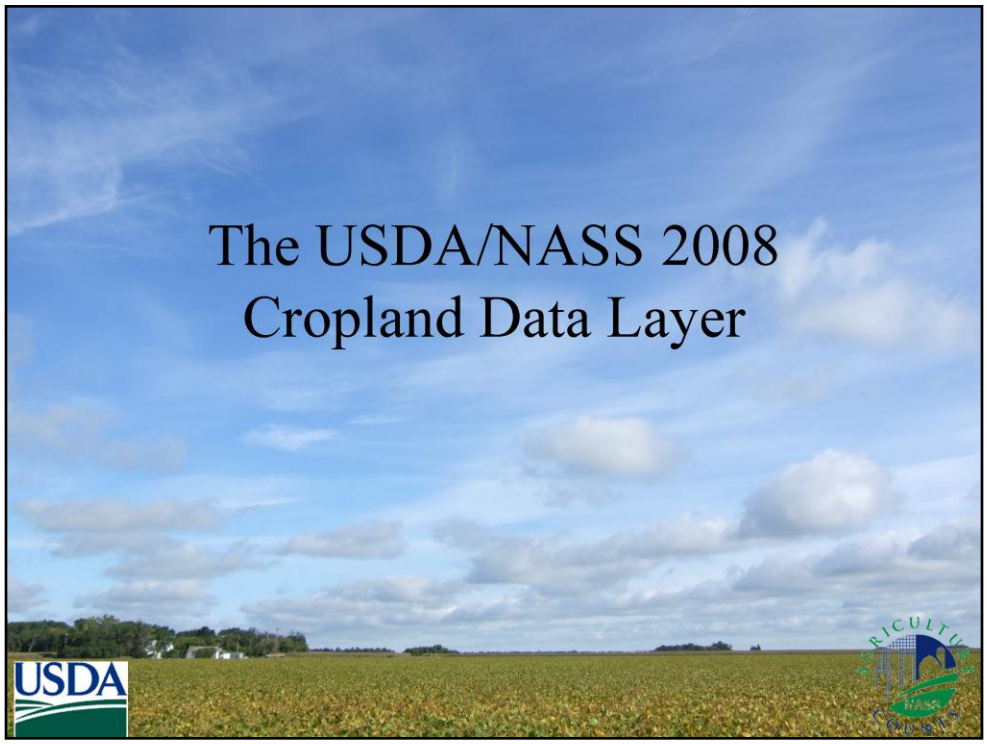


This is a brief overview of the NASS remote sensing activities and methodology involving acreage estimation and the Cropland Data Layer.

Cropland Data Layer (CDL) Defined

- "Census by Satellite"
- Geospatial (map registered) image product
- Depicts accurate field crop locations
- Crop targeted – unique, timely, detailed land cover classification
- Robust because of large inventory of ground truth & satellite imagery

•The CDL program began in earnest in 1997 with the ability to deliver geospatial content annually to customers who were interested in annual crop land cover updates. Prior to the creation of the CDL product, estimates were provided in tabular format, with pictures/outputs depicting the results.

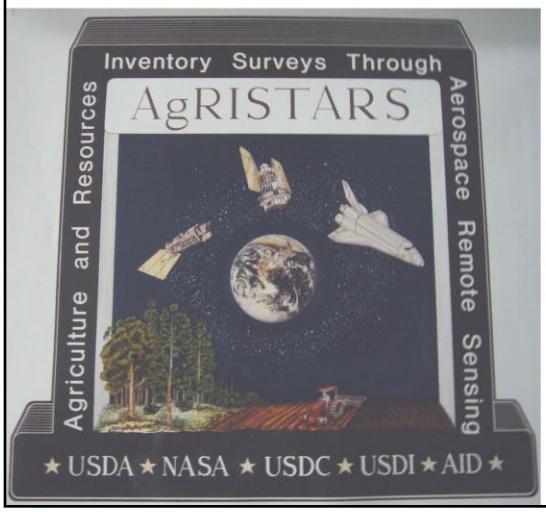
•The CDL can be considered a "Census by Satellite", as it is a comprehensive land use classification covering an entire state, and uses ortho-rectified imagery, to accurately locate and identify field crops.

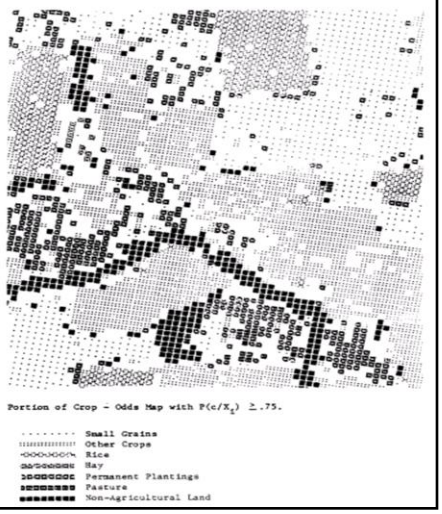
•The CDL is now produced in-season to produce operational timely estimates for decision maker support.

•The CDL utilizes a comprehensive and robust archive of AWiFS satellite imagery from the Foreign Ag Service along with ground truth data provided by the Farm Service Agency.

Cropland Data Layer (CDL) History

- Legacy program
  - Issues: Budget/Satellites/Technology





•The Cropland Data Layer Program is a legacy program within NASS that has undergone much refinement since it's inception. However, the program has grown immensely these past few years and has overcome issues such as constrained budgets, failing satellites, and technological innovations.

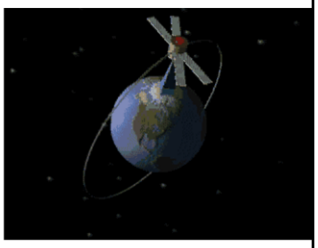
•PEDITOR, the original NASS image classifier, was originally written in the 1970's, and was updated and maintained since by NASS. It was developed during the early 70's using Purdue University's LARSYS system as a basis for further development. NASS and the University of Illinois Center for Advanced Computing developed a customized program called EDITOR. It ported to other computer platforms by NASS and the name modified to PEDITOR.

•NASS has supported PEDITOR throughout the LACIE and AgRISTARS programs and continued until 2006, as PEDITOR was updated and modified to run on the latest desktop platforms utilizing some of the original algorithms from the LARSYS project. However, in 2005 alternative software application testing began to improve upon PEDITOR's success.

Cropland Data Layer (CDL) Discussion

Operational Paradigm

- Deliver in-season acreage estimates
  - Multiple times during growing season
- Increase program scope/coverage
  - "Manifest Destiny"
- Public domain crop specific
  - Land Cover Classification
  - Creation of derivative analysis products



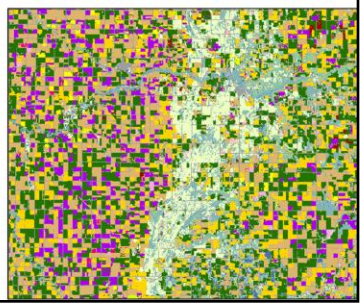
•The Cropland Data Layer (CDL) is now operational providing in-season estimates for decision support in our NASS Field Offices and Agricultural Statistics Board. Estimates are delivered multiple times during the growing season, helping improve agency estimates.

•The CDL program strives to cover all NASS speculative program crops of Corn, Soybeans, Wheat, and Cotton in crop year 2009, providing improved acreage estimates throughout the growing season as more farmer reported and satellite data are utilized.

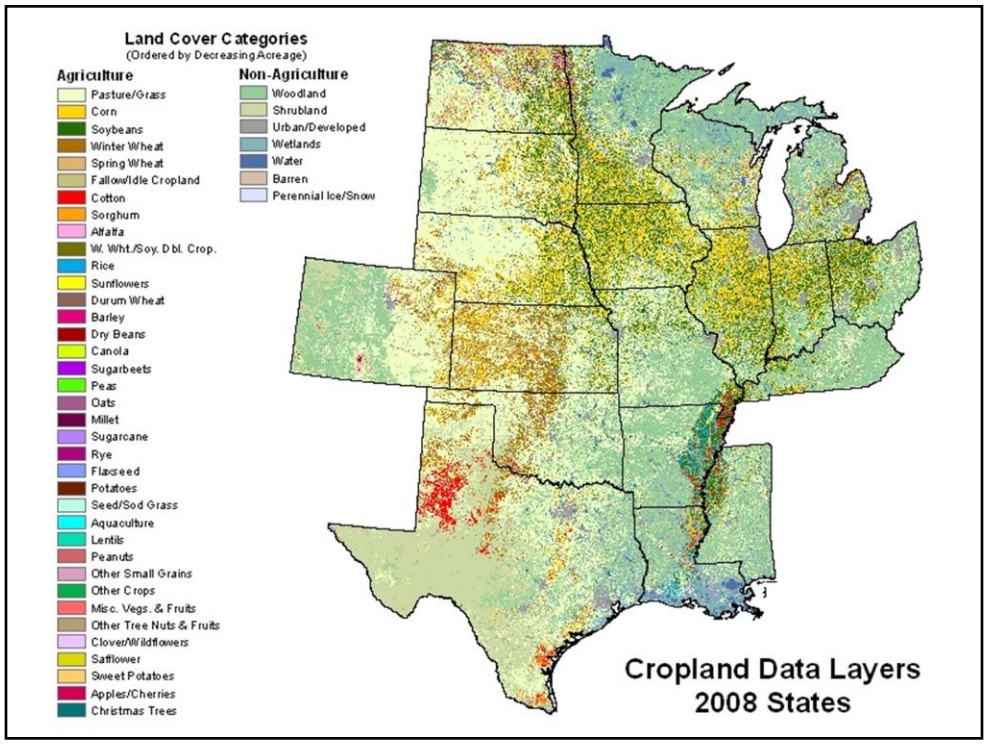
•The CDL is a publically releasable crop specific land cover classification that focuses primarily on mapping cultivated fields and providing an update on the agriculture landscape. Additionally, crop intensity products will be released in 2009.

CDL Program Objectives

- **Provide timely, accurate, useful indications**
  - Measurable error
  - Unbiased/independent estimator
  - State, County, Agricultural Statistics Districts
- **Operationalize indications delivery**
  - For June, August, and October
    - Agricultural Statistics Board
    - Field Offices
  - Update planted area
- **Output crop specific CDL**
  - Distribute to public at the cost of reproduction
    - NRCS Geospatial Data Gateway

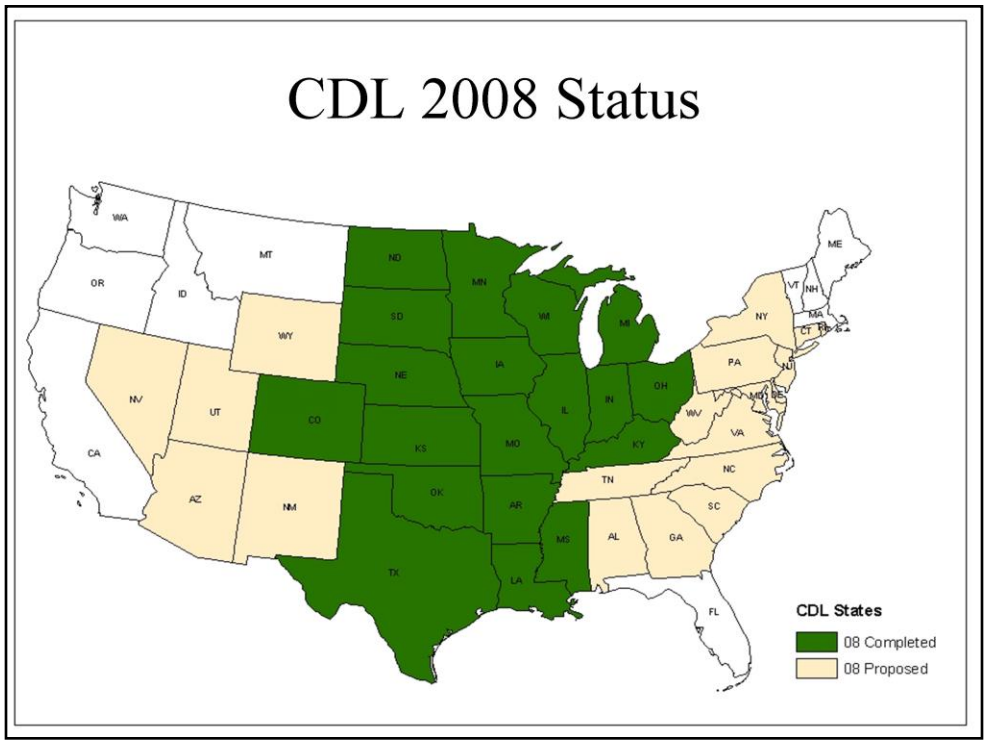


•The Cropland Data Layer Program provides internal state/district/county level indications of major commodities with accuracy and variance statistics, and secondarily provides the public with "statewide“ categorized output products.
•Satellite based estimates will not completely replace surveys of farmers for several reasons: acreage planted estimates are completed in June, before the crop canopy fully develops on summer crops and cloud problems can cause loss of large areas.
•This project builds upon the June Ag Survey and improves upon the enumerator collected ground survey data with satellite imagery to create an unbiased statistical estimator of crop area. Remote sensing provides an update on planted acreage as the growing season progresses, effectively lowering the crop estimate C.V.
•The CDL is available for free for download from the NRCS Geospatial Data Gateway,

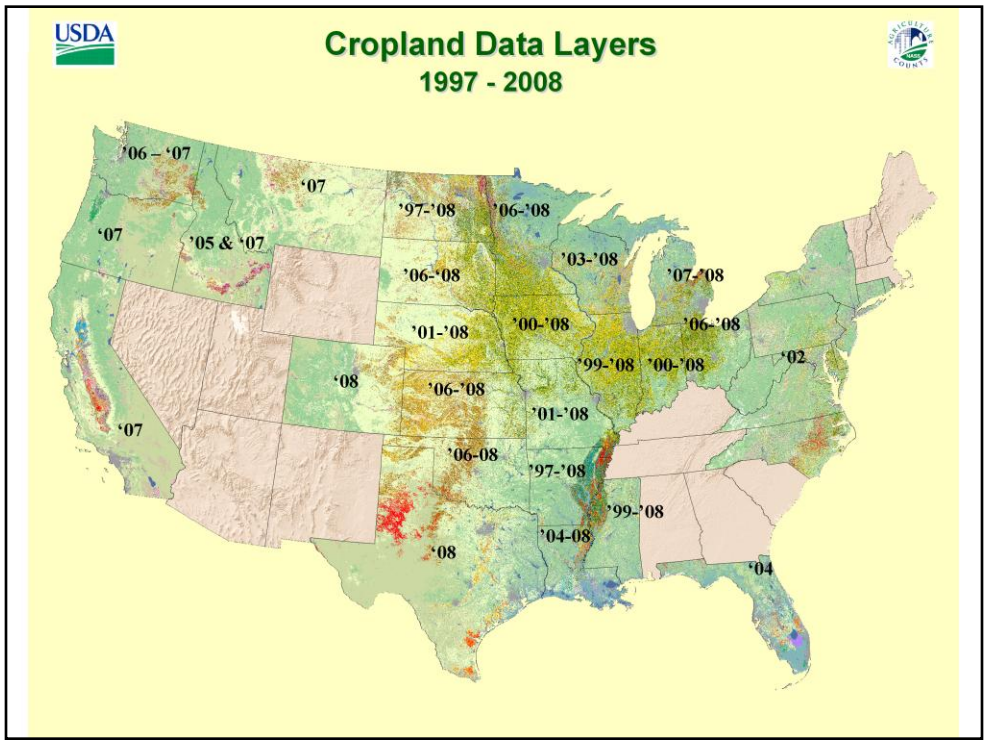


The 2008 CDL released in March 2009.

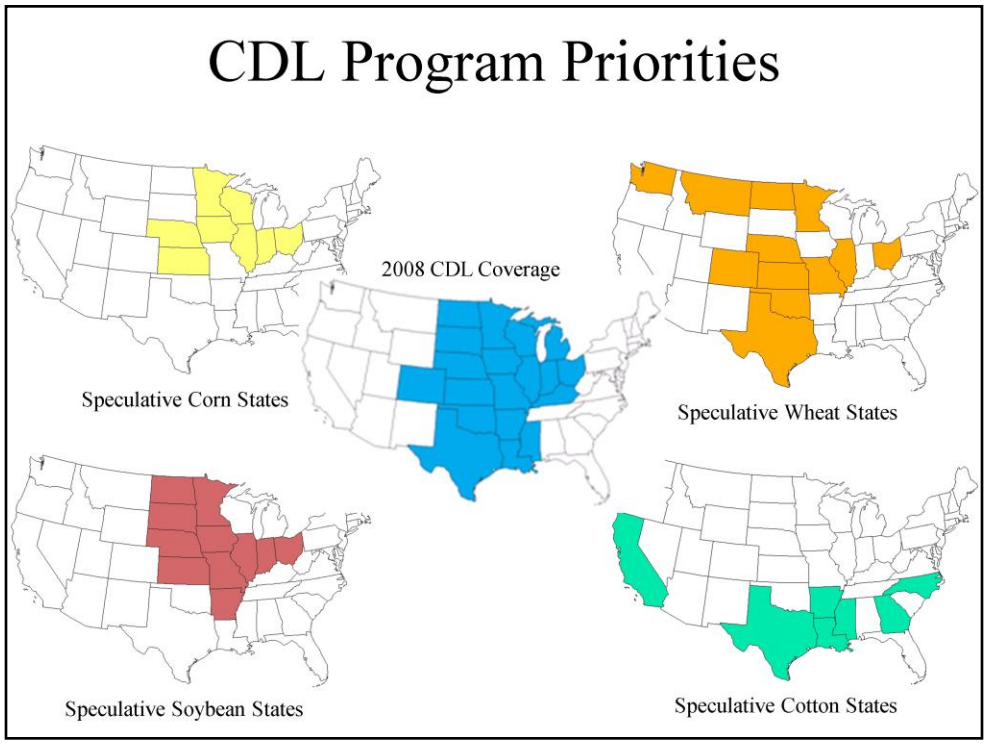
No farmer reported data is revealed, nor can it be derived in the publicly releasable Cropland Data Layer product.



The 2008 green colored states are completed with efforts ongoing to process many of the tan shaded states for summer 2009 release.



This graphic shows the CDL historical coverage and the year's in production. Numerous partnerships were brokered to cover additional states.

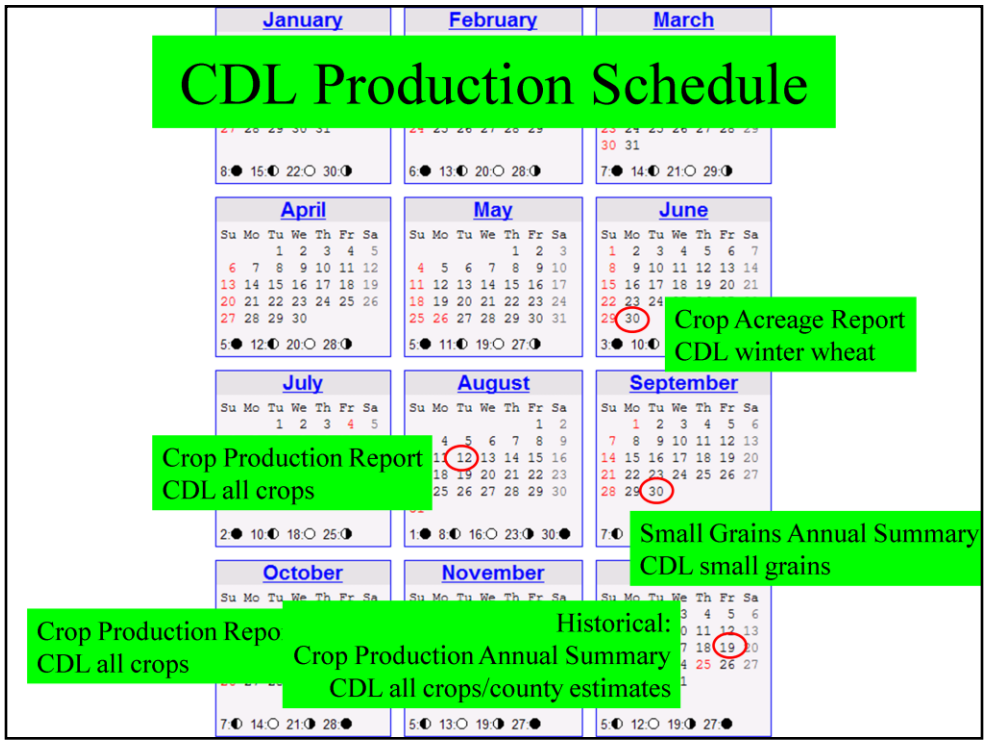


These are the designated NASS Speculative States for Corn, Wheat, Soybean and Cotton. Note the CDL is covering all but a few wheat and cotton states, which will be covered in 2009.

2008 CDL Coverage

Commodity	CDL States	US Total Acres (mill)	% US Total
Corn	18	78,177	92
Soybeans	18	74,374	91
Rice	5	2,924	82
Wheat	13	40,252	70
Cotton	4	7,755	66
Potatoes	11	1,058	34

This table shows the processed 2008 CDL states by commodity along with the total amount of US acreage. The last column represents the total coverage of the CDL program as a percentage of the whole US.



•The CDL program has undergone major restructuring and modernization these past few years.

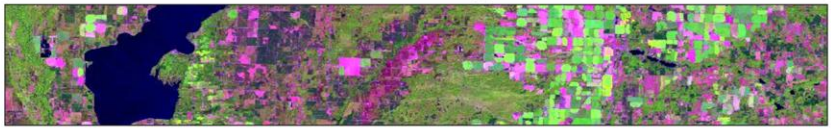
•The new efficiencies allow for in-season crop acreage estimates, that were not previously possible with our older methods. The historical PEDITOR method delivered state and county level indications in late December for the Crop Production Annual Summary.

•The CDL is now able to deliver state/district/county estimates throughout the season starting with Winter Wheat for the June 30th Crop Acreage Report.

•The next deadline is the August Crop Production Report where major corn and soybean district level estimates are produced for our Agency stakeholders.

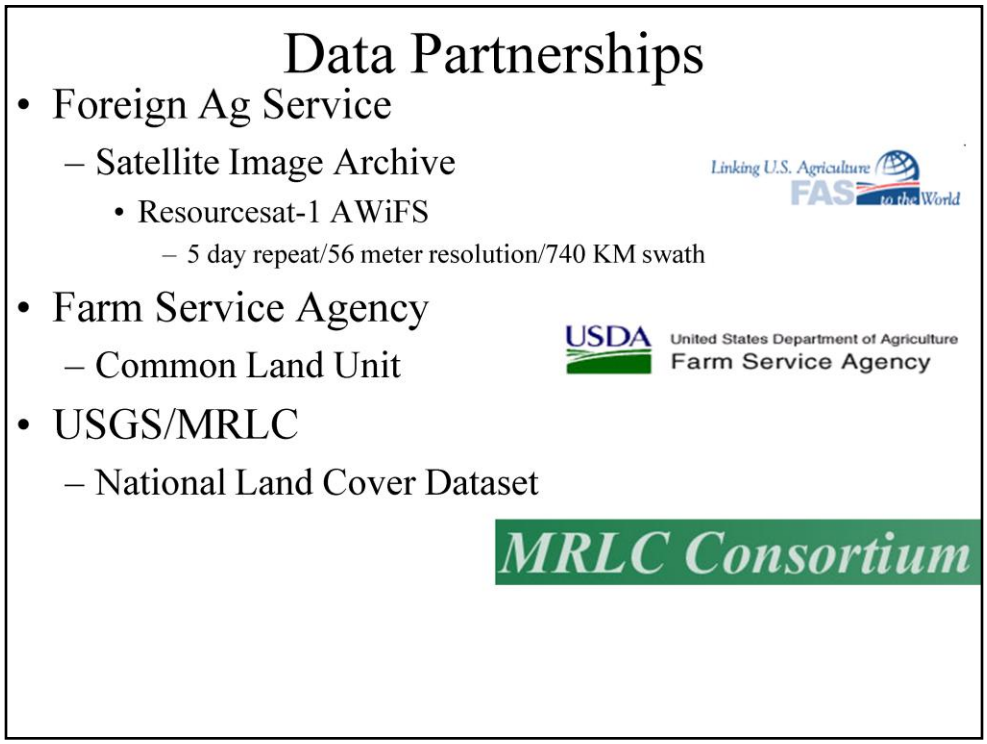
•The small grains estimates are produced for the September Small Grains Annual Summary and all CDL states are updated for the production of end-of-season estimates for the October Crop Production Report.

CDL Program



- Inputs
  - Resourcesat-1 AWiFS imagery
  - Farm Service Agency – Common Land Unit
  - JAS segment boundaries & summaries
  - Ancillary data
- Outputs
  - Acreage Estimates
  - Cropland Data Layer

These are the inputs/outputs of the CDL program.



Partnerships are vital to the CDL Program. The Foreign Ag Service/Satellite Image Archive has provided imagery to NASS through a MOU since 1997. Currently, the Resourcesat-1 AWiFS satellite is operational in USDA. The Farm Service Agency/Common Land Unit has been in production for the CDL program since 2006, providing robust ground truth for supervised classifications. Since 2006, the USGS/Multi-Resolution Land Characteristics Consortium/National Land Cover Dataset has been utilized to improve classification accuracy in the non-ag domain.

IRS Resourcesat-1 AWiFS Imagery

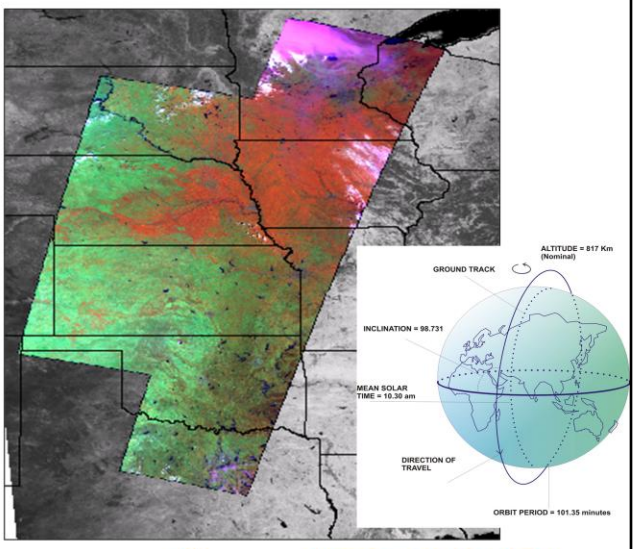
340 km swath per head
740 km combined

5-day revisit

4 spectral bands
- B2: 0.52 - 0.59
- B3: 0.62 - 0.68
- B4: 0.76 – 0.86
- B5: 1.55 – 1.7

56 m nadir/70 m field edges

Data provided by EOTEC Corporation

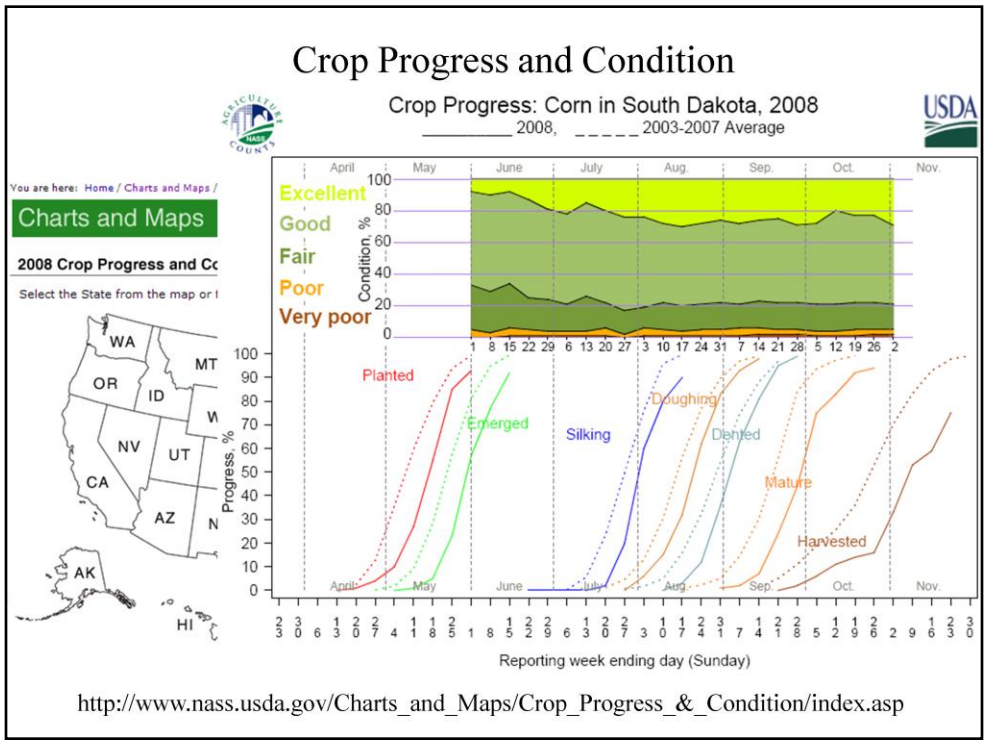


EOTec

13 Aug 2007

Department of Space
Indian Space Research Organisation

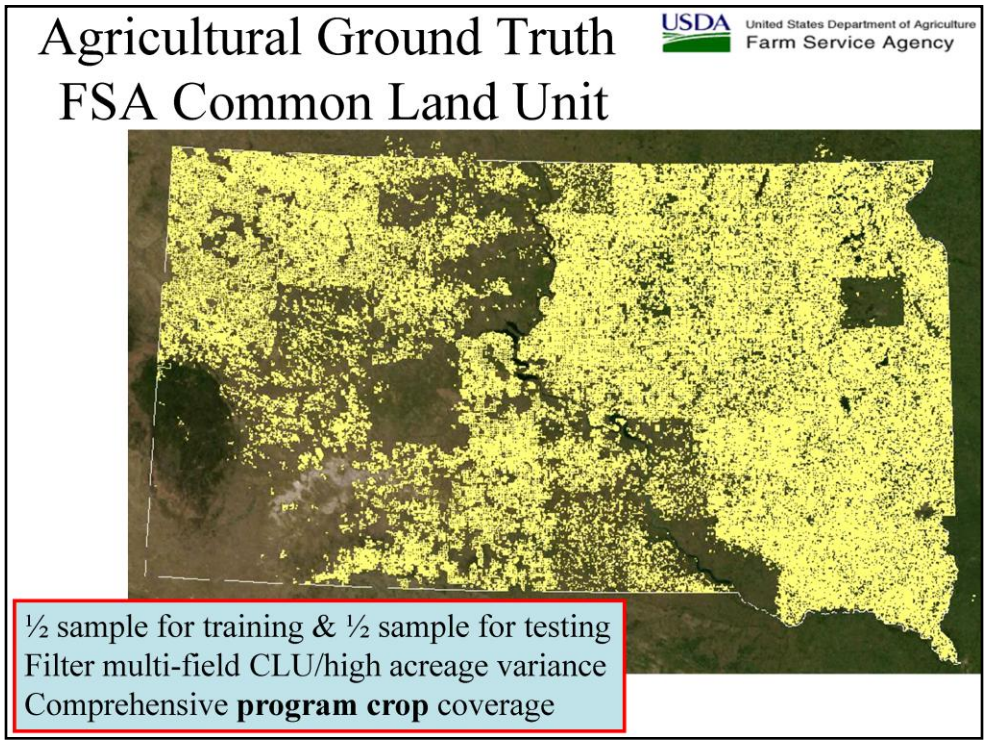
IRS – Indian Remote Sensing Resourcesat-1 Advanced Wide Field Sensor. Launched October 2003, large swath width, 5 day repeat coverage, with 24 day orbital repeat, 4 spectral bands and 56 meters resolution. The imagery is purchased ortho-rectified by USDA and is ingested in 16 bit format. Note the large swath width is capable of covering large Midwestern States in one overpass.



Crop Progress and Condition graphical products are shown throughout the key stages of the crops phenological cycle. The progress of the crop through each stage is shown as a percentage in the lower half of the graphic. Condition Ratings are shown in the upper part of the graphic. This product is generated in all states and will include the major commodities reported in each state. Separate pages are embedded in the .PDF file, in alphabetical order, for each commodity. This product is updated around midweek throughout the growing season.

http://www.nass.usda.gov/Charts_and_Maps/Crop_Progress_&_Condition/index.asp

This product helps NASS imagery analysts differentiate each crops' phenology and help them with their planning on which scenes to select for the CDL classification process.



•The Common Land Unit (CLU) is quite a robust ground truth dataset. There are comprehensive coverage of program crops from farmer signups. FSA is more comprehensive and less labor intensive at the cost on not being a true probability-based sample.

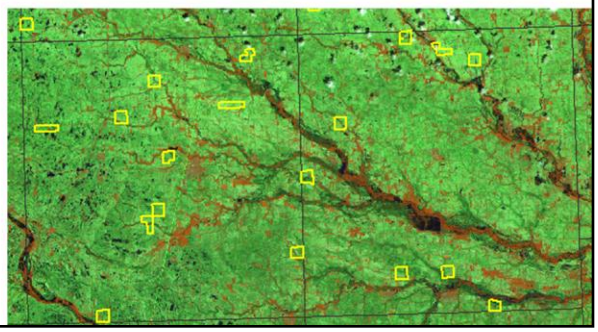
•The CDL splits the CLU into two separate datasets, one half is used for training the classifier, while the other is used for testing/accuracy assessment validation. Providing two independent datasets.

•The CDL uses filtering on the CLU's to prime it for remote sensing usage. There are instances with CLU crop fields that have a many to one relationship with only one polygon (i.e., a corn and soybean field within one polygon), and it is difficult to separate out which field is which, so that CLU is dropped from consideration. Other instances occur when the reported acreage does not match the digitized polygon.

- The FSA data are very comprehensive but have a bias toward "program" crops. The FSA data are digitized in FSA county field offices.
- Ground truth data must be map projected and rasterized to the same characteristics as imagery data.
- Before deriving training signatures the ground truth polygons are buffered inward so as not to use edge pixels for training. A distance of 56 - 84 meters using the mid-resolution satellite imagery has been found to work best.

NASS June Ag Survey

- Probability based
- Area frame stratification based on land use
- Sample units one square mile

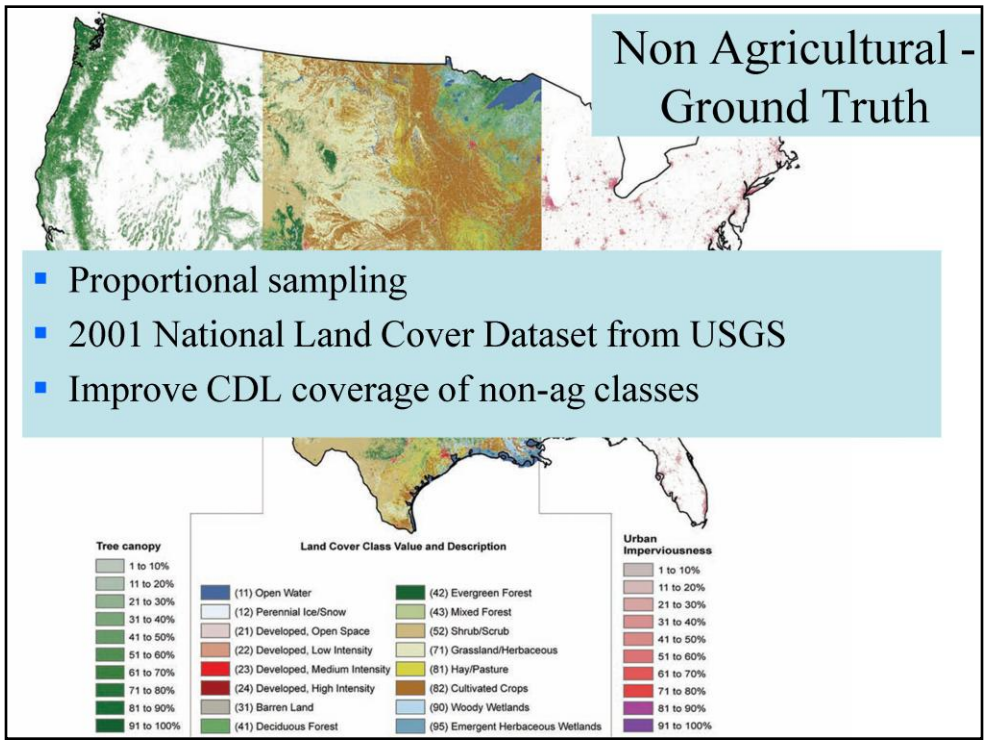


•Every June approximately 41,000 farms are visited by enumerators as part of the USDA/NASS June Agricultural Survey (JAS). These farmers are asked to report the acreage, by crop, that has been planted or that they intend to plant, and the acreage they expect to harvest. Approximately 11,000 area segments are selected nationwide for the JAS. This represents approximately 2.5 percent of the total land area in the entire United States.

•The segment size can range in size from four to eight square miles in open range areas to about 1 square mile in cultivated areas to 0.1 of a square mile in urban areas. This division allows intensively cultivated land segments to be selected with a greater frequency than those in less intensively cultivated areas. Sample segments representing cultivated areas are selected at a rate of about 1 out of 125, whereas sample segments in land use classifications with decreasing amounts of cultivated land are selected at rates ranging from 1 out of 250 to 1 out of 500.

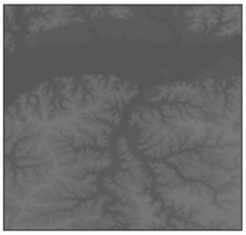
•The JAS data are statistically robust because they are based on a probability survey. The 150 – 400 square miles of ground truth collected during the JAS provides a basis for building the regression model estimate.

•Every field/land use within each segment is accounted for on the survey.



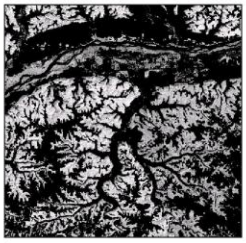
The USGS/NLCD 2001 product is utilized for improving the non-ag domain. The non-ag areas are sampled at the same rate as FSA data to maintain appropriate ground truth proportions.

Ancillary Data – USGS/NASA Products



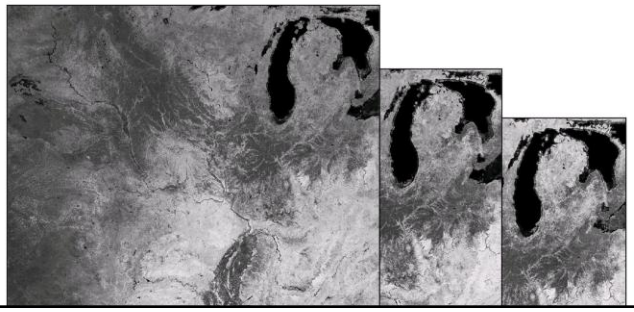
Elevation

Imperviousness



Forest Canopy

NASA MODIS Terra
(16-day NDVI composite)



The USGS National Elevation Dataset along with the percent imperviousness and forest canopy products were used to help separate the non-ag domain. Additionally, NASA's MODIS 16 day NDVI composites are used to help identify winter wheat fall green-up, as USDA did not have fall AWiFS collections in fall 2007. MODIS is also beneficial where there is only sparse AWiFS coverage.

Commercial Software Suite

- Imagery Preparation
  - ERDAS Imagine

- Image classification
  - Decision tree software
    - See5.0 www.rulequest.com

- Ground Truth Preparation
  - ESRI ArcGIS

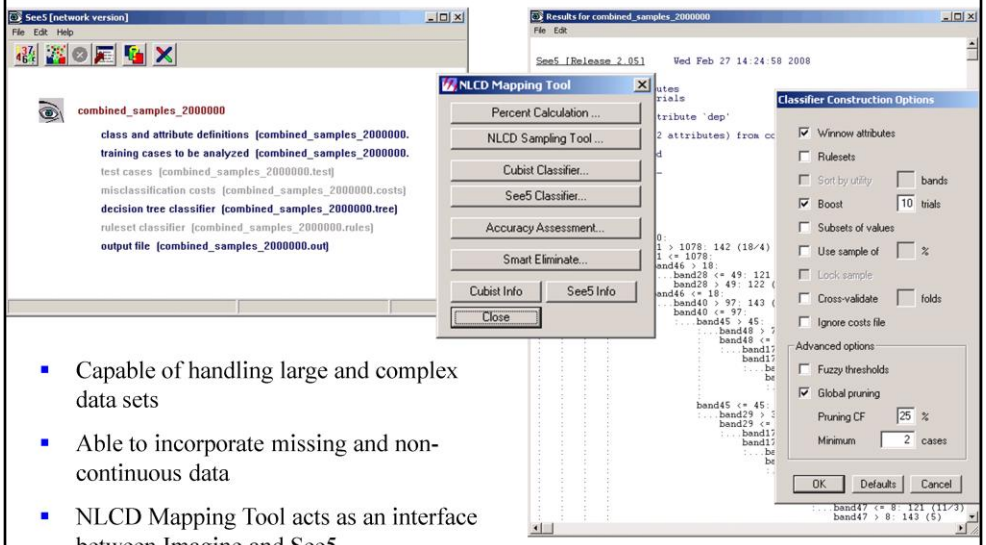
- Acreage Estimation
  - SAS/IML workshop

•The CDL Program has moved to commercial software packages to produce the CDL products. These software include: Leica Geosystems ERDAS Imagine 9.1 for imagery preparation, See5.0, decision tree software, to perform the classification and ESRI ArcGIS 9.3 to prepare the ground truth.

•SAS IML workshop is producing state/district/county estimates.

•The NLCD Extension was developed by USGS; integrated with ERDAS Imagine and acts as an interface between See5 and Imagine.

Classification – See5 Decision Tree



- Capable of handling large and complex data sets
- Able to incorporate missing and non-continuous data
- NLCD Mapping Tool acts as an interface between Imagine and See5

- Decision Tree software has become increasingly popular within the remote sensing community for a number of reasons. It is non-parametric by nature and thus not reliant on the assumption of the input data being normally distributed. Second, it is efficient to construct and capable of handling large and complex data sets. Third, it is able to incorporate missing and non-continuous data. Specifically, Rulequest See5.0 commercial software is commonly utilized because an extension was written by USGS to easily interface with ERDAS Imagine image processing software and it incorporates an advanced “boosting” classification tree algorithm known to improve outcomes further.
- As with the traditional maximum likelihood method, decision trees are also a “supervised” classification technique and thus reliant on statistically representative ground truth information.

Some good classification tree references:

Friedl and Brodley, 1997 Decision Tree Classification of Land Cover from Remotely Sensed Imagery, RSE

DeFries and Cheugn-Wai Chan, 2000 Multiple Criteria for Evaluation Machine Learning Algorithms for Land Cover Classification from Satellite Data, RSE

Lawrence and Wright, 2001 Rule-Based Classification Systems Using Classification and Regression Tree (CART) Analysis

Bricklemyer et al., 2005 Predicting tillage practices and agricultural soil disturbance in north central Montana with Landsat imagery

Quinlan, 2006 Bagging, Boosting, and C4.5

Accuracy Statistics

Crop-specific covers only	*Correct Pixels	Accuracy	Error	Kappa
OVERALL ACCURACY	2306428	87.51%	12.49%	0.8416

Cover Type	Attribute Code	*Correct Pixels	Producer's Accuracy	Omission Error	Kappa	User's Accuracy	Commission Error	Cond'l Kappa
Corn	1	803251	94.29%	5.71%	0.9342	95.78%	4.22%	0.9513
Sorghum	4	9047	46.40%	53.60%	0.4630	79.16%	20.84%	0.7909
Soybeans	5	707383	95.03%	4.97%	0.9439	97.72%	2.28%	0.9741
Sunflowers	6	107195	85.99%	14.01%	0.8572	92.15%	7.85%	0.9199
Sweet corn	12	0	0.00%	100.00%	0.0000	n/a	n/a	n/a
Popcorn	13	627	64.77%	35.23%	0.6477	94.86%	5.14%	0.9486
Barley	21	1995	25.85%	74.15%	0.2582	64.17%	35.83%	0.6412
Durum wheat	22	280	13.53%	86.47%	0.1352	57.49%	42.51%	0.5748
Spring wheat	23	255912	86.02%	13.98%	0.8537	91.04%	8.96%	0.9060
Winter wheat	24	310316	84.53%	15.47%	0.8368	94.00%	6.00%	0.9363
Other grains	25	92	4.75%	95.25%	0.0475	64.79%	35.21%	0.6478
WW / Soybeans	26	10	3.66%	96.34%	0.0366	100.00%	0.00%	1.0000
Rye	27	126	6.71%	93.29%	0.0671	78.26%	21.74%	0.7825
Oats	28	2799	14.85%	85.15%	0.1479	58.23%	41.77%	0.5810
Millet	29	12879	49.50%	50.50%	0.4936	74.76%	25.24%	0.7465
Flaxseed	32	150	17.69%	82.31%	0.1769	66.37%	33.63%	0.6637
Safflower	33	212	14.89%	85.11%	0.1488	57.30%	42.70%	0.5729
Rape seed	34	0	0.00%	100.00%	0.0000	n/a	n/a	n/a
Alfalfa	36	56603	56.37%	43.63%	0.5593	90.69%	9.31%	0.9054
Beets	41	14	8.86%	91.14%	0.0886	93.33%	6.67%	0.9333
Dry beans	42	827	51.02%	48.98%	0.5101	94.19%	5.81%	0.9419
Other crops	44	8	13.33%	86.67%	0.1333	42.11%	57.89%	0.4210
Misc. vegetables	47	0	0.00%	100.00%	0.0000	n/a	n/a	n/a
Watermelon	48	0	n/a	n/a	n/a	0.00%	100.00%	0.0000
Lentils	52	253	87.54%	12.46%	0.8754	99.61%	0.39%	0.9961
Peas	53	950	35.26%	64.74%	0.3525	88.29%	11.71%	0.8828
Herbs	57	639	78.21%	21.79%	0.7821	98.61%	1.39%	0.9861
Clover / Wildflowers	58	27	13.24%	86.76%	0.1323	93.10%	6.90%	0.9310
Seed / Sod Grass	59	319	18.07%	81.93%	0.1807	89.86%	10.14%	0.8986
Idle / Fallow	61	34514	56.97%	43.03%	0.5668	82.73%	17.27%	0.8257
Apples	68	0	0.00%	100.00%	0.0000	n/a	n/a	n/a

*Correct Pixels represents the total number of independent validation pixels correctly identifed in the error matrix.

Accuracy statistics such as this are published in the metadata. Note the high accuracies in the FSA large area program crops. The CDL strives to obtain accuracies in the 90's for the major program crops.

Producer's Accuracy: relates to the probability that a ground truth pixel will be correctly mapped and measures errors of omission.

Errors of Omission: occur when a pixel is excluded from the correct category

User's Accuracy: indicates the probability that a pixel from the classification actually matches the ground truth data and measures errors of commission

Errors of Commission: occur when a pixel is included in an incorrect category

Kappa Coefficient: A statistics measure of agreement, beyond chance, between two maps (e.g. output map of classification and ground truth map.)

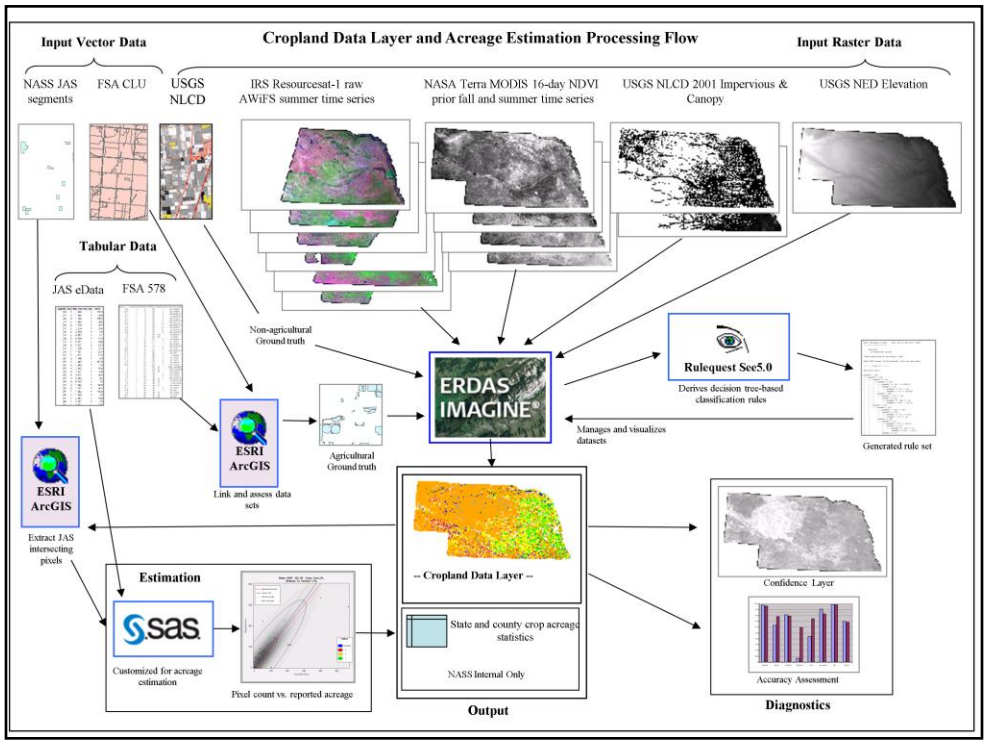
CDL Metadata

- Published on each CDL product

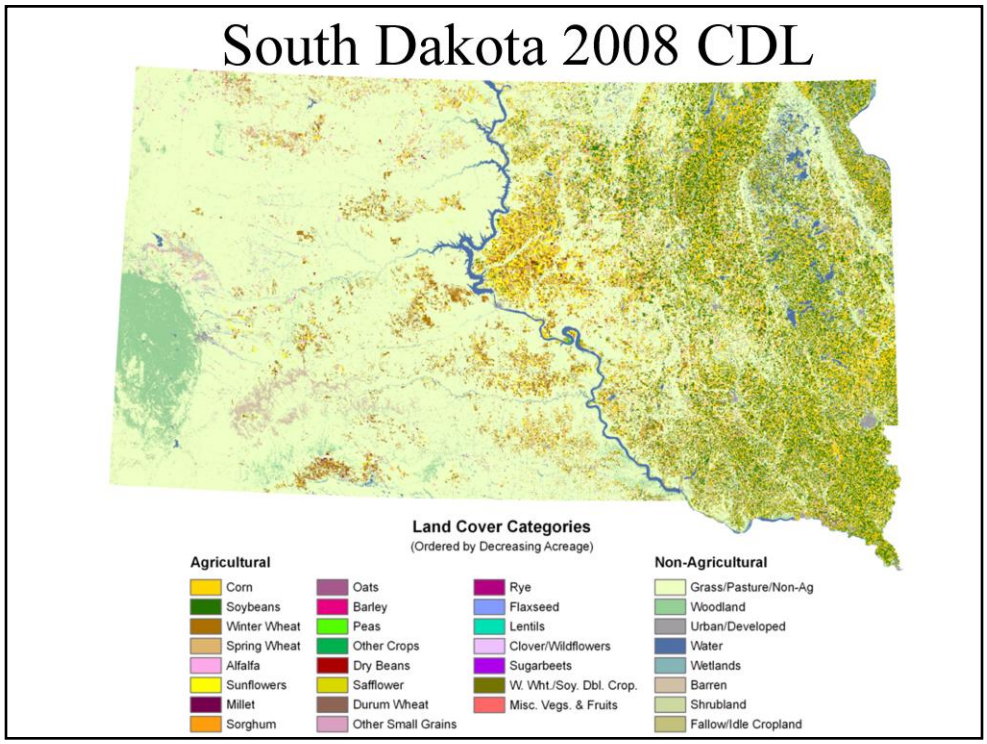
```
Raster
Attribute Domain Values and Definitions: ROW CROPS 1-20

Categorization Code   Land Cover
        "1"           Corn
        "2"           Cotton
        "3"           Rice
        "4"           Sorghum
        "5"           Soybeans
        "6"           Sunflowers
        "10"          Peanuts
        "11"          Tobacco
        "12"          Sweet Corn
        "13"          Popcorn or Ornamental Corn


Map_Projection_Name: Albers Conical Equal Area
Albers_Conical_Equal_Area:
Standard_Parallel: 29.500000
Standard_Parallel: 45.500000
Longitude_of_Central_Meridian: -96.000000
Latitude_of_Projection_Origin: 23.000000
False_Easting: 0.000000
False_Northing: 0.000000
Planar_Coordinate_Information:
Planar_Coordinate_Encoding_Method: row and column
Coordinate_Representation:
Abscissa_Resolution: 56
Ordinate_Resolution: 56
Planar_Distance_Units: meters
Geodetic_Model:
Horizontal_Datum_Name: North American Datum of 1983
Ellipsoid_Name: Geodetic Reference System 80
Semi-major_Axis: 6378137.000000
Denominator_of_Flattening_Ratio: 298.257223563
```

```
CLASSIFICATION INPUTS:
AWIFS DATE 20080413 PATH 264 ROW(S)&QUADRANT(S) 35b 40d 45bd
AWIFS DATE 20080418 PATH 265 ROW(S)&QUADRANT(S) 35bd 40abcd 45a
AWIFS DATE 20080427 PATH 262 ROW(S)&QUADRANT(S) 40bd
AWIFS DATE 20080428 PATH 267 ROW(S)&QUADRANT(S) 40d 45bd
AWIFS DATE 20080503 PATH 268 ROW(S)&QUADRANT(S) 35bd 40bcd 45ab
AWIFS DATE 20080512 PATH 265 ROW(S)&QUADRANT(S) 40bcd 45abd
AWIFS DATE 20080517 PATH 266 ROW(S)&QUADRANT(S) 35d 40bd 45b
AWIFS DATE 20080606 PATH 270 ROW(S)&QUADRANT(S) 40d 45b
AWIFS DATE 20080614 PATH 262 ROW(S)&QUADRANT(S) 35bd 40bd 45b
AWIFS DATE 20080625 PATH 269 ROW(S)&QUADRANT(S) 40d 45b 50bd
AWIFS DATE 20080629 PATH 265 ROW(S)&QUADRANT(S) 40bd 45b
AWIFS DATE 20080704 PATH 266 ROW(S)&QUADRANT(S) 35a 40d 45bd
AWIFS DATE 20080713 PATH 263 ROW(S)&QUADRANT(S) 35abcd 40abd 45
AWIFS DATE 20080715 PATH 273 ROW(S)&QUADRANT(S) 35cd 40abcd 45a
AWIFS DATE 20080802 PATH 267 ROW(S)&QUADRANT(S) 35d 40abcd 45ab
AWIFS DATE 20080808 PATH 273 ROW(S)&QUADRANT(S) 35d 40bc 45a
AWIFS DATE 20080812 PATH 269 ROW(S)&QUADRANT(S) 35c 40ac 45a
AWIFS DATE 20080904 PATH 264 ROW(S)&QUADRANT(S) 40bd 45bd
AWIFS DATE 20080909 PATH 265 ROW(S)&QUADRANT(S) 35bd 40bd
AWIFS DATE 20080914 PATH 266 ROW(S)&QUADRANT(S) 40d 45bd
AWIFS DATE 20080915 PATH 271 ROW(S)&QUADRANT(S) 45bd 50b

MODIS 16 DAY NDVI COMPOSITE DATE 20071016
MODIS 16 DAY NDVI COMPOSITE DATE 20071101
MODIS 16 DAY NDVI COMPOSITE DATE 20071117
MODIS 16 DAY NDVI COMPOSITE DATE 20080305
MODIS 16 DAY NDVI COMPOSITE DATE 20080321
MODIS 16 DAY NDVI COMPOSITE DATE 20080406
MODIS 16 DAY NDVI COMPOSITE DATE 20080422
MODIS 16 DAY NDVI COMPOSITE DATE 20080508
MODIS 16 DAY NDVI COMPOSITE DATE 20080524
MODIS 16 DAY NDVI COMPOSITE DATE 20080609

USGS, NATIONAL ELEVATION DATASET ELEVATION
USGS, NATIONAL LAND COVER DATASET 2001 TREE CANOPY
USGS, NATIONAL LAND COVER DATASET 2001 IMPERVIOUSNESS
```

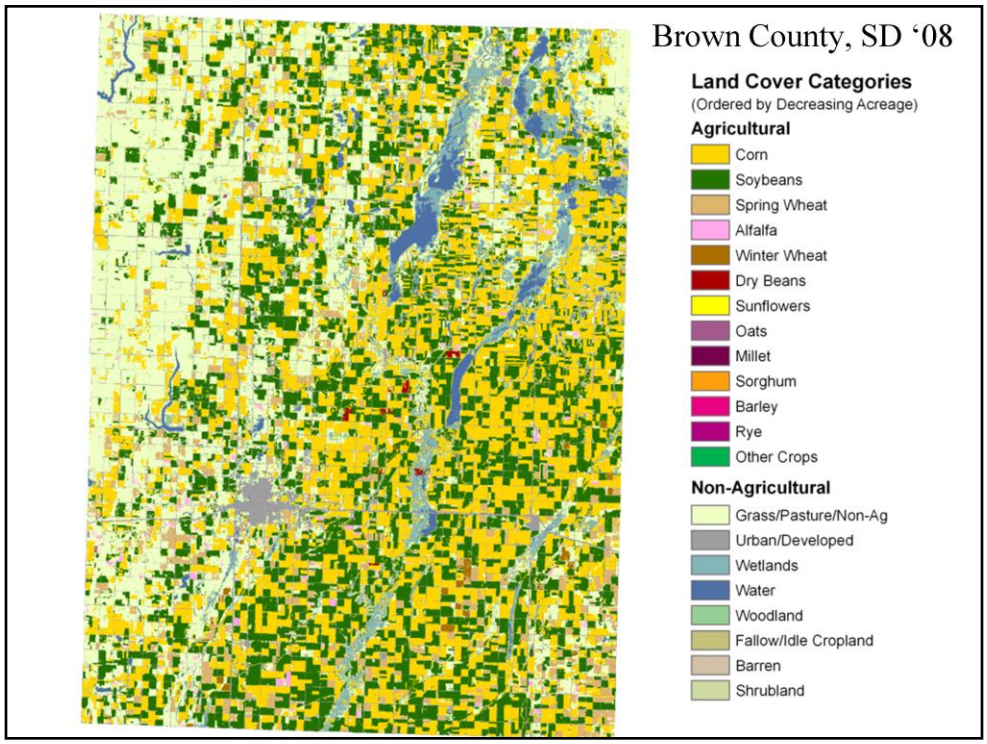
Standardized Metadata is published on each CDL product.



The NASS CDL production flowchart.

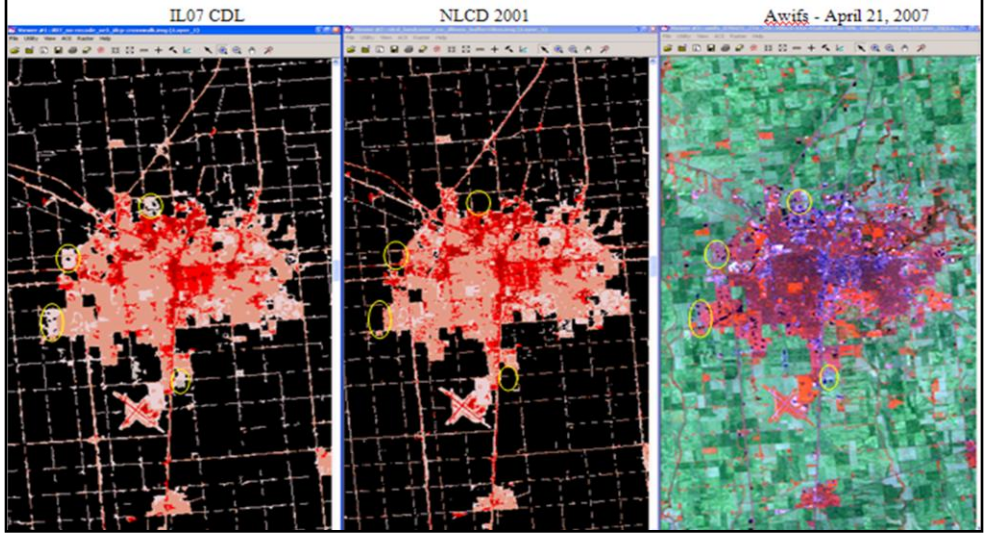


This is the finalized 2008 CDL for South Dakota



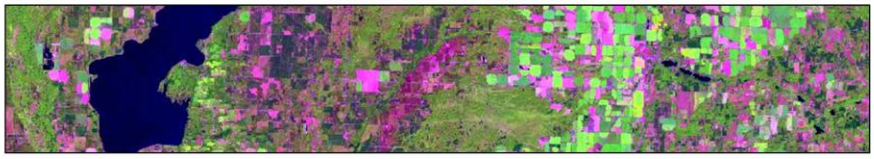
Enlargement of 2008 Brown County, SD CDL

Non Ag NLCD Updates (urban sprawl)

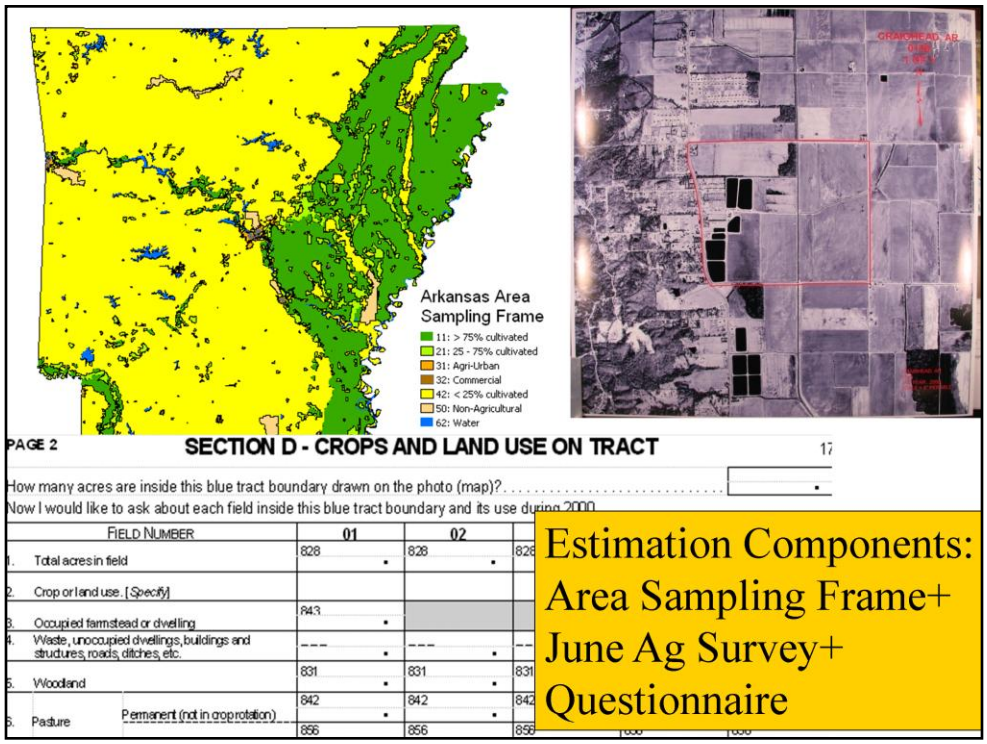


This example shows the USGS/NLCD 2001 product of Champaign, Illinois in the middle of the graphic. The far right graphic shows a raw AWiFS satellite scene taken on 4.21.2007. The far left scene is the CDL of 2007, note the CDL is catching farmland to urban conversion in the yellow circled areas on the city perimeter.

Remote Sensing
Regression Estimation



The following slides show acreage based regression estimation



•The three components of acreage estimation: the June Ag Survey area component, with questionnaire and an Area Sampling Frame segment.

•The Area Sampling Frame (ASF) is a stratification of each state into broad land use categories according to the percentage of cropland present. Since 1978, satellite imagery has been the major input into stratification of land based on broad land cover definitions. Previously, aerial photography mosaics were used. Each year NASS replaces some of the area frames because the land use changes over time.

•The ASF is stratified using visual interpretation of satellite imagery. This has led to improved statistical precision of numerous area frame-based estimates, including coverage estimates for major probability surveys and the Census of Agriculture. In addition, beginning in 1978 and continuing today, area sampling frames have been converted from paper-based products, subject to fire and loss, to digital versions which are more accurate and better protected from loss.

•The sampling frames are constructed by defining blocks of land whose boundaries are physical features on the ground (roads, railroads, rivers, etc.). These blocks of land cover the entire state, do not overlap, and are placed in strata based on the percent of land in the block that is cultivated. The strata allow for efficient sampling of the land, as an agriculturally intensive area will be more heavily sampled than a non ag intensive area.

•Every June, approximately 41,000 farms are visited by enumerators as part of the June Agricultural Survey.

•The unit of observation is the tract, which may contain one or more fields or land uses and represents a particular land operator's acreage within a segment. The enumerators draws off field boundaries onto the National Aerial Photography Program's (NAPP) 1:8,000 scale black and white aerial photos where the segment is located, according to their observations and the farmer reported information. The fields are labeled and the cover type is recorded using a grease pencil on the aerial photo.

•Enumerators account for every field/land use type within a segment. They assign each field a cover type based upon a fixed set of land use classes for each state. Every field within a segment must fit into one of the pre-defined classes.

•This is a sample of the questionnaire from which the enumerator asks the farmer for information.

•Enumerators record the grower's responses on cover type and acreage for each field in a segment on the JAS questionnaire. The questionnaire is directly linked to the NAPP 1:8,000 segment photo by referencing the field number between the questionnaire and the photo.

•The farmer reported data is only used internal to NASS and cannot be derived from the public output Cropland Data Layer. Farmer reported data is held strictly confidential by NASS to calculate aggregated statistics.

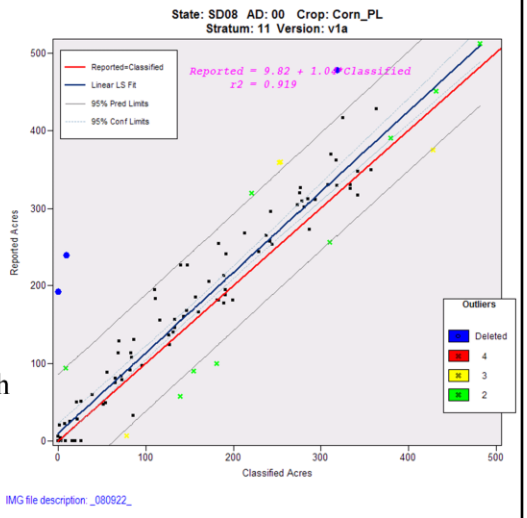
Regression-based Acreage Estimator

Regression used to relate categorized pixel counts to the ground reference data

- **(X) – Cropland Data Layer (CDL) classified acres**
- **(Y) – June Agricultural Survey (JAS) reported acres**

Using both CDL and JAS acreage results in estimates with reduced error rates over JAS alone

Outlier segment detection - removal from regression analysis

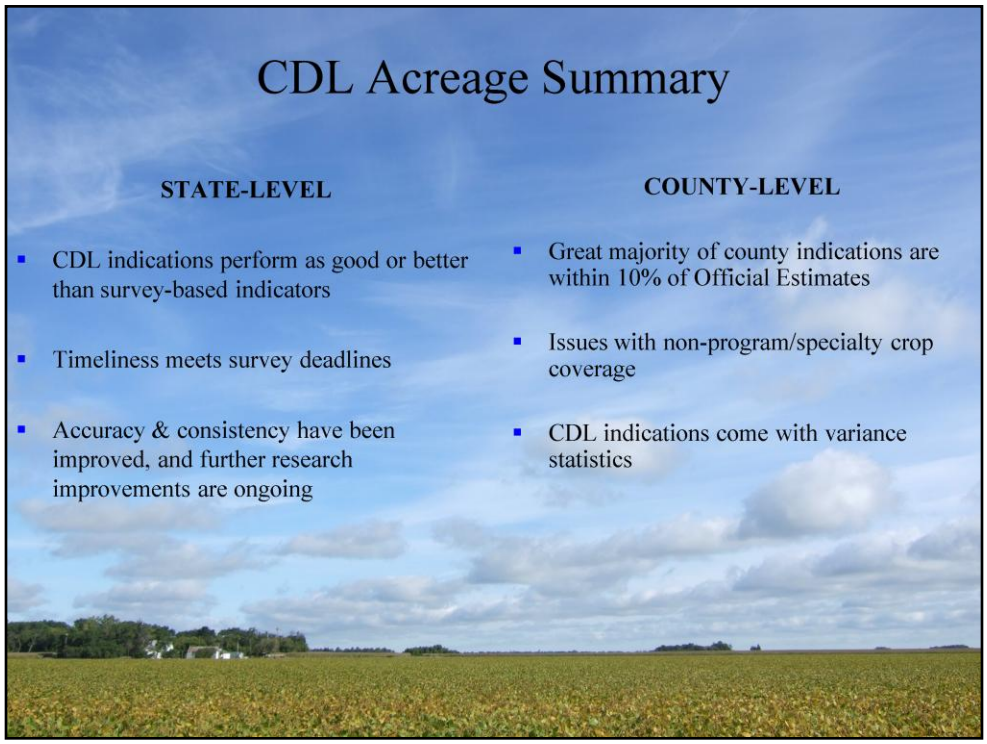


Acreage not just about counting pixels

A simple linear regression is performed at the segment level on the June Ag survey segments and classified pixel data.

- Y =a + bx (formula)
- Y = Dependent variable: June Agricultural Survey reported acres
- X = Independent variable Cropland Data Layer (Remote Sensing) classified acres
- b = slope
- a = Intercept

- Where available, regression is chosen as the preferred type of estimation. This approach essentially corrects the area sample (ground only) estimate based on the relationship found between reported data and classified pixels in each stratum where it is used.
- Regression adjusts the direct expansion estimate based on pixel information. It usually leads to an estimate with a much lower variance than direct expansion alone.
- Segments, called outliers, which do not fit the linear relationship estimated by the regression are reviewed; if errors are found, that segment may be removed from consideration in the regression analysis.
- This graph shows the approximately linear relationship between corn acres reported during the ground survey and acres classified to corn in the process of producing a CDL.
- Several possible outliers are visible.



CDL Summary…

•The CDL is now distributed via the NRCS Geospatial Data Gateway.